

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2002, 06:43:24 ; Search time 38.27 seconds
(without alignments)
1076.834 Million cell updates/sec

Title: US-09-652-292-2
Perfect score: 2765
Sequence: 1 MGHSPVPLCLASVSLGGL.....CHRONSTGIPYSRIEISAAS 541

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:.*
1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	626.5	22.7	457	2 E70070	metabolite transpo
2	595	21.5	461	2 D70073	metabolite transpo
3	594.5	21.5	580	2 D86426	hypothetical prote
4	567	20.5	521	2 G84864	probable membrane
5	545	19.7	580	2 D84772	probable sugar tra
6	541	19.6	582	2 F71431	hypothetical prote
7	521.5	18.9	433	2 G86812	D-xylose proton-sy
8	520.5	18.8	557	2 T38125	myo-inositol trans
9	508	18.4	493	2 A85433	sugar transporter
10	499	18.0	464	2 F65079	galactose-proton s
11	497.5	18.0	547	2 A84442	membrane transport
12	497	18.0	511	2 A84537	probable sugar tra
13	497	18.0	606	2 T27072	hypothetical prote
14	496	17.9	464	2 F85951	galactose-proton s
15	495	17.9	472	2 B26430	L-arabinose isomer
16	495	17.9	472	2 B85936	hypothetical prote
17	495	17.9	560	2 T51485	sugar transporter-
18	491	17.8	523	2 S25015	monosaccharide tra
19	489.5	17.7	472	2 S47089	arabinose-proton s
20	488.5	17.7	464	2 F69587	L-arabinose transp
21	488.5	17.7	522	2 B86246	glucose transporte
22	484.5	17.5	511	2 H84536	probable sugar tra
23	484.5	17.5	522	2 S12042	glucose transport
24	483	17.5	516	2 T12199	monosaccharid tran
25	482	17.4	508	2 T05156	probable glucose t
26	480.5	17.4	549	2 T14606	probable sugar tra
27	473.5	17.1	575	2 T43400	myo-inositol trans
28	471.5	17.1	639	2 T23658	hypothetical prote
29	467	16.9	523	2 T10122	hexose transport p

30	466.5	16.9	508	2 G84564	probable sugar tra
31	464.5	16.8	468	2 S10014	glucose transport
32	460	16.6	487	2 E96782	hypothetical prote
33	460	16.6	491	2 A26430	xylose transport p
34	460	16.6	491	2 B86096	xylose-proton symp
35	455.5	16.5	502	2 B70845	probable sygar tra
36	452.5	16.4	482	2 B69803	metabolite transpo
37	451.5	16.3	613	2 T27077	hypothetical prote
38	450	16.3	473	2 G89789	sugar transporter
39	450	16.3	547	2 C84593	probable sugar tra
40	438	15.8	584	2 S69555	myo-inositol trans
41	435	15.7	490	2 T14545	probable sugar tra
42	434.5	15.7	513	2 T01506	probable hexose tr
43	430.5	15.6	522	2 T10150	sugar transport pr
44	426.5	15.4	504	2 T00450	probable monosacch
45	426	15.4	472	2 T35662	probable sugar tra

ALIGNMENTS

RESULT 1

E70070

metabolite transport protein homolog ywtG - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: E70070

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabet, C.; Ferrari,
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gal
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil
A:Reference number: A69580; MUID:98044033
A:Accession: E70070

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-457 <KUN>

A:Cross-references: GB:Z99122; GB:AL009126; NID:92636029; PIDN:CAB15600.1; PID:g26361

A:Experimental source: strain 168

C:Genetics:

A:Gene: ywtG

C:Superfamily: glucose transport protein

Query Match 22.7%; Score 626.5; DB 2; Length 457;
Best Local Similarity 30.1%; Pred. No. 2.1e-34;
Matches 160; Conservative 85; Mismatches 179; Indels 107; Gaps 8;

Oy 17 LGLTFGYELAVISGALLPLQDFGLSCLEQEPFLVGLSLGALLASLVGLFDYCYGRKO 76

Db * 15 LGLALYGYDTGVISGAILPMKELGLNATEGLVSSSLVGLGAILGSAAGKGLTDRGRKK 74

Oy 77 ATILGNLVLAGSLTGLAGSLAWLVGLRAVVGFAISLSSMACCIYVSELVGPQRGVLY 136

Db 75 ATMAAALFCIGGLGVALPNTGVNLFRIILGLAVGTSTTIVPLVLSLAPKHKRGALS 134

Oy 137 SYEAGITVGIILLSTALYNALACTPWGHRMFNGWATAPAVLQSLSLPLP-----AGT 189

Db 135 SUNQLMITVGIILSVIVNYIFADAE-AWRWMLGLAAVPSLLLLIGILFMPESPRLWLTNG 193

Oy 190 DETATHKDLIPLOQ-----GEAPKLGPRPRYSFGLDIFRARDNMRGTTVGLGL 238

Db 194 EESKAKKILEKLRGTYKDIOEIHDIKEAKQDEG----GLKELFDPW--VRPALIAGLGL 247

QY 239 VLFOQLTGQPNVLCYASTIFSSVGFHGGSSAVLASVGLGAVKVAATLTAMGLVDRAGRA 298
 Db 248 AFLQOFIGNITIIYAPKFTNVP- GNSASILGTGTGVTGVTNVLMTLVAIKIDIGRKP 306
 QY 299 LILACALMALSVSGIGLVSAFVPMDSGSPCLAVPNATGOTGLPGDGLLODSSLPPPR 358
 Db 307 LLLFGNAGVSLIVALVNLFF----- 329
 QY 359 TNEDQREPLSTAKTKPHRSGDPSAPPRALSSALPCPPLPARGHALLRWATALLCMV 418
 Db 330 -----NWTPAA-----SWTTVICLGV 345
 QY 419 FVSASFSGFPGVTWLVSEIYPVEIRGRAFAFCNSFNMAANLFISSLFDLIGTIGLSWT 478
 Db 346 FIVFAVSGWGPVWVMLPELPHVRGIGTGVSTLMLHVGVLIVSLYPPILMEAGISYL 405
 QY 479 FLVGLTAVLGLGFTLYLPVETKQSLAEIDQOQKRRFTLSFGHRSNGT 529
 Db 406 FLIVAAIGIMAFVRFKVTETKRSLEETEQDLDRKNGGGAAGKQOTVG 456
 RESULT 2
 D70073
 metabolite transport protein homolog yxcC - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
 C:Accession: D70073
 R:Kunze, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertelli, C.; Bron, S.; Brouillet, S.; Bruschini, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinol, Y., M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumsteln, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033
 A:Accession: D70073
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-461 <KUN>
 A:Cross-references: GB:Z99124; GB:AL009126; NID:g2636442; PIDN:CAB16017.1; PID:g2636527
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yxcC
 C:Superfamily: glucose transport protein

Db 194 GSEEEARRIMNITHDPKDIEMELAEAKQGEAEK-----KETTGLVKKAK-WIRPMLLIG 246
 QY 236 LGLVLFQOLTGQPNVLCYASTIFSSVGFHGGSSAVLASVGLGAVKVAATLTAMGLVDRAG 295
 Db 247 VGLAIFOQAVGINTVIYAPKFTNVP- GTSASALGTGTGVTNVLMTLVAIKIDIGRKP 305
 QY 296 RRALLAGCALMALSVSGIGLVSAFVPMDSGSPCLAVPNATGOTGLPGDGLLODSSLPP 355
 Db 306 RKLLIWSGVITLSLAAL-----SGVLLTGLS- 334
 QY 356 IPRNEDQREPLSTAKTKPHRSGDPSAPPRALSSALPCPPLPARGHALLRWATALLC 415
 Db 335 -----ASTA-----WMTVVF 344
 QY 416 LMVFAFSFGFPGVTWLVSEIYPVEIRGRAFAFCNSFNMAANLFISSLFDLIGTIGL 475
 Db 345 LGVIVFYQATWGPVWVMLPELPHVRGIGTGVSTLMLHVGVLIVSLYPPILMEAGISYL 404
 QY 476 SWTFLYGLTAVLGLGFTLYLPVETKQSLAEIDQOQKRRFTLSFGHRSNGT 529
 Db 405 AMVFMVSVICLISFFFAFYMPETKGSLEIEIASLKKR 444
 RESULT 3
 D86426
 hypothetical protein AAG30955.1 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: D86426
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alor Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dewar, L.; ansen, N.F.; Hughes, B.; Hultzar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maffi, R.; Marzia Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: AB6141; MUID:21016719
 A:Accession: D86426
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-580 <STO>
 A:Cross-references: GB:AE005172; NID:g11120774; PIDN:AAG30955.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 21.5%; Score 594.5; DB 2; Length 580;
 Best Local Similarity 29.6%; Pred. No. 3.7e-32;
 Matches 173; Conservative 94; Mismatches 217; Indels 101; Gaps 16;

QY 4 SPVPLICASVSLGLTGFYELAVISGALLPLQDF---GLSCLEQEFVLGSLIGALL 60
 Db 25 NPVYLRFAFSAG-IGGLFGYDTGTVISGALLYIRDFKSVDRNTWLOEMIVSNVAGAI 83
 QY 61 ASLVGGFLIDCYGRKQALICNSLVLLAGSLTIGLAGSLAWLVLGRAVVGFAISLSSMACC 120
 Db 84 GAAIGWANDKLGRRSAILMADFLLGATIMAAFPNLSLLVYGRVFGVGLGVGNASTAP 143
 QY 121 IYVSELVGPQRGVLSVLEAGITVGLLSYALNYALAGTPWGRHMFHWGATAPAVLQSL 180
 Db 144 LYISEASPAKIRCALVSTNGFLITGQGLSYLINLAFTDVTGTWRMLGTAGIPALLQV 203
 QY 181 SLFLP-----ACTDETA-----THKDLIP---LQGEAPKLP 211
 Db 204 LMFTLPSRWLYRKGRREEAKAILRRIYSAEDVEQEIIRALKDSVETEIEGSSSEKI-- 261
 QY 212 GRPRYSFLDLFRARDNMRCRTTTLGLVLFQOLTGQPNVLCYASTIFSSVGFHGGSSAVL 271
 Db 262 -----NNIKLCAKTVRG-LTAGVGLVQFQGVGINTVWYISPTIVQLAGFASNRALL 315

QY 272 ASVGLGAVKVAATLTAMGLVDGRAGRAILLACALMALSVSGIGLVFAVPMDSGSPCLA 331
 Db 316 USLVTAGLNAGFIIISIFIDIRGRKKLLI-----ISLFGV-IISUGI----- 357
 QY 332 VFNATG-----QTGLPGDSGL-----LQSSSLPPIPTNEDQREPILSTAKTKPK----- 376
 Db 358 ---LTGVFEAATHAPALSSLETORFNILSCPDYKSAmntNAWDCMTCLKASPSGCGYCS 414
 QY 377 -----HPRS-----GDSAPP-----RIALSSALPGPPLPARGHALLRTALLCLMVFVS 421
 Db 415 SPIGKEHPCACWISDSDVKDCHNENRLMYTRGCP-----SNFGWFALLGLGLYII 465
 QY 422 AFSPGCGPVTWLVSEIYVPEIRGRAFAFCNSFNAAANLISLSFLDLIGTIGLSTWTFLL 481
 Db 466 FSPGNGVTPWVNSEIITPLRFRGICGGIAATANNISNLIVAQSFSLFEATIGTSMTEFLI 525
 QY 482 YGLTAVLGLGFIYLPVETPKGOSLAIEDQOQKRRFTLSFGHRQN 526
 Db 526 FGVISVIALLFVWVCPETKGMPEIEKMLERRSMERKFKKKS 570

RESULT 4
 G84864
 probable membrane transporter [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: G84864
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: AB4420; MUID:20083487
 A:Accession: G84864
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-521 <STO>
 A:Cross-references: GB:AE002093; NID:g2289003; PIDN:AB864332.1; GSPDB:GN00139
 A:Gene: At2g43330
 A:Map position: 2
 C:Superfamily: glucose transport protein

Query Match 20.5%; Score 567; DB 2; Length 521;
 Best Local Similarity 27.8%; Pred. No. 2.2e-30;
 Matches 150; Conservative 85; Mismatches 183; Indels 122; Gaps 6;

QY 17 LGGLTFGYELAVISGALLPLQLDL-----GLSCLEQEFVLGSLIGALLA 61
 Db 40 ICGLLFGDTGTVISGALLYIKDDFEVVKOSFLQVNVSSFTSSKLETIVSMALVGAMTG 99
 QY 62 SLVGGFLIDCYGRKQAILGSLNLVLAGSLTGLAGSLAWLVGRVAVGPAISLSSMACCI 121
 Db 100 AAGGWINDYQYGRKATLADYVFAAGAVMAAPDPVYLISGLRLVGLGVCAVAPV 159
 QY 122 YVSELVGRQVRLVSLYEAGITVIGLISYALNYALAGTPGWRHMFGNATAPVQLQSL 181
 Db 160 YIAEASPEVRGGLVSTNVLMTITGGQFLSYLNSAFTQVPGTWRWMLGVSVPVAVIQFTL 219
 QY 182 LLFLP-----ACTDETATHKDIPLQGEAPKLGPRPRYSFLDLF 222
 Db 220 MLFMPESPRLWPKMKRKAIEIQLVATYDISREIDEIDHLSAAEEKEKRTKTVGLDVF 279
 QY 223 RARDNMRGRTTGLGLVLFQQLTGOPNVLCYASTIFSSVGFHGGSAVLASVGLGAVKA 282
 Db 280 RSKE-LRLAFLAGAGLQAFQQTGINTVMYYSPTIVOMAGFHSNQLALFLSLIVAAANNA 338
 QY 283 ATLTAMGLVDGRAGRAILLAG-----CALMALSVSGIGLVFAVPMDSGSPCLAVPNATG 337
 Db 339 GTVVGIVYFDHCGRKKLALLSSLFGVILSLILSVSFF-----K 376
 QY 338 QTGLPGDSQLQDSSLPIPTNEDQREPILSTAKTKPKPRSGDPSAPPRLLSALPG 397

Db 377 QSETSDGG----- 395
 QY 398 PPLPARGHALLRTALLCLMVFVSFAFGPVTWLVSEIYVPEIRGRAFAFCNSFNMA 457
 Db 386 -----LYGWLAVLGLALYVFPAPGMPVPMVNSEIYPOQYRGICGGNSATVNM 436
 QY 458 ANLFTLSFLDLIGTIGLSTWTFLLYGLTAVLGLGFIYLPVETPKGOSLAIEDQOQKRRF 517
 Db 437 SNLVAQTETLIAEAAGTGTMTFLILAGIAVLAVIFVFPETQGLTFSEVEQIKERAY 496

RESULT 5
 D84772
 probable sugar transporter [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: D84772
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: AB4420; MUID:20083487
 A:Accession: D84772
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-580 <STO>
 A:Cross-references: GB:AE002093; NID:g4263781; PIDN:AAD15441.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g35740
 A:Map position: 2

Query Match 19.7%; Score 545; DB 2; Length 580;
 Best Local Similarity 28.6%; Pred. No. 7e-29;
 Matches 164; Conservative 99; Mismatches 205; Indels 106; Gaps 18;

QY 4 SPVPLPLCASVSLGLLTFGYELAVISGALLPLQLDLFG---LSCLEQEFVLGSLIGALL 60
 Db 23 TPVIMRLALSAG-IGGLFGYNTGTAGALLYIKKEFGEVDNKTWLIQIIVSMTVAGAIV 81
 QY 61 ASLVGFLIDCYGRKQAILGSLNLVLAGSLTGLAGSLAWLVGRVAVGPAISLSSMACC 120
 Db 82 GAAGGWINDYQYGRKQAILGSLNLVLAGSLTGLAGSLAWLVGRVAVGPAISLSSMACC 141
 QY 121 YVSELVGRQVRLVSLYEAGITVIGLISYALNYALAGTPGWRHMFGNATAPVQLQSL 180
 Db 142 LYISEMSPARIGALVSTNGLITGGQFLSYLNLAFVHTPGTWRWMLGVSIPAIIQFC 201
 QY 181 SLFLPAG-----TDETATHKDI-----PLOGGEA-----PKLGPGRP 214
 Db 202 LMLTLPESPRLWPKMKRKAIEIQLVATYDISREIDEIDHLSAAEEKEKRTKTVGLDVF 257
 QY 215 RYSELDLFR-ARDN--MRGRTTGLGLVLFQQLTGOPNVLCYASTIFSSVGFHGGSSAV- 270
 Db 258 -HTFSKDLRGALSNPVVRHGLAGITVQVAAQPGVINTVMYYSPTILOPAGYASNTAMA 316
 QY 271 --LASVGLGAVKVAATLTAMGLVDGRAGRAILLACALMALSVSGIGLVFAVPMDSGSP 328
 Db 317 LALITSLGNV---GSVSMFMFVDYGRKKLM-----ISMFGII-----T 354
 QY 329 CLAVPNATGOTGLPGDGLLQDSSLPIPTNEDQRE-----PILS---TAKTKPH-- 377
 Db 355 CLVILAAVNEA-----SNHAPKIDK--RDRSNFAKNATCPAFAPPTASRSPSNW 403
 QY 378 -----PRSGDPSAPPRLLSALPGPPLPARGHALLR-----WTALLCL 416
 Db 404 NCMKCIQYDCGFCNGAQEYAPGACIVQSDMKALCHSKGRTEFFKDCGCKSKFYLAIVFL 463
 QY 417 MVFVSASFSGFGPVTWLVSEIYVPEIRGRAFAFCNSFNAAANLISLSFLDLIGTIGL 476
 Db 464 GLYIIVAFMGITVPMVNSEIYPLRYGLAGGIAAVSNMNSLVSEITFLTITNAVGS 523

QY 477 WTFLYGLTAVLGLGFIYLFVFPETKQSLAEIDQ 510
 Db 524 GTFLFAGSAGVGLFTFLVLPETKGLQFEVEK 557

RESULT 6
 F71431
 Hypothetical protein - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 A:Variety: columbia
 C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
 C:Accession: F71431
 R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel
 vanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
 Nature 391, 485-488, 1998
 A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
 erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
 C.; Chalwatzis, N.
 A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
 A:Reference number: A71400; MUID:98121113
 A:Accession: F71431
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-582 <BEV>
 A:Cross-references: GB:297341; NID:g22445004
 C:Genetics:
 A:Map position: 4COP9-4C3845
 C:Superfamily: glucose transport protein

Query Match 19.68; Score 541; DB 2; Length 582;
 Best Local Similarity 28.88; Pred. No. 1.3e-28;
 Matches 161; Conservative 91; Mismatches 231; Indels 76; Gaps 13;
 QY 4 SPPVLPCASVSLGGTTFYGLAVISGALLPLQLDF---GLSCLEQEFVIGSLIGALL 60
 Db 24 TPYIMRLALSAG-IGGLFGYDGTGVSIGALLFFKEDEVDKWTWOSTIVMAVAGAI 82
 QY 61 ASLVGGFLIDCYRKQAILGSLNVLGSLTLGLAGSLAWLVLGKAVVGFPAISLSSMACC 120
 Db 83 GAAGVGNIDKFRMSIADVLEITGAVMAFAPAPVPIIVGRIFVGVGMASTSP 142
 QY 121 IYSELVGPQRGVLSLYEAGITVGLISYALNYALAGTPWGRHMFQWATAPVQLSL 180
 Db 143 LYISEASPARIGALYSTNGLITGQFFSYLINLAFVHTPTGTRWMLGVAGVPAIVQFV 202
 QY 181 SLFLPAG-----TDETATHKDLIPLOGEAPK-----LGPGR 213
 Db 203 LMLSLPSPRWLYRKORIAESRAILERIYPADVEAEMALKL-SVEAEKDEAIIIG--- 258
 QY 214 PRYSFLDLFRAR-----DNRGRRTVGLGLVLPQGLVFOQLGQPNVLCYASTIFSSVGFHGG 266
 Db 259 -----DSFSAKLKGFAGNPVVRVLAAGITVQVAAQFVGINVMYSPSIVQFAGYASN 312
 QY 267 SSALVASVGLGAVKVAATLMTAGLVDRAGRALL-----LAGCALMALSVSGILVSFA 320
 Db 313 KTAMALLSLITSLGNALGSIVSMFVDYGRKMLIISFGITACLIILATVPSQAII-HA 371
 QY 321 VPMDSGSCSLAVPNATGOTGLPGDGLGSLDSSLPPTPTNEDQREPILSTAKTKPHPRS 380
 Db 372 PKIDAFESTFAPNATCSAYP-----LAAENAPP-----SRNCKMKLRSECGFCAS 419
 QY 381 G-DPSAPPLALSSALPGPPLPARGHALLR-----WTALLCLMVFSVAFSGFGVPT 431
 Db 420 GVQYAPGACVVLSDMKATCSSRGRTFFKQCPKSKGFGLAIVFLGLIYVYAFQMGVTP 479
 QY 432 WLVLSEIYVPEIRGRAFAFNSFWAANLFTSLFDLIGTIGLSWTFLLYGLTAVLGLG 491
 Db 480 WIVNSEIYPLVRGLGGGIAVSNWVNLIVSEFSLTHALGSGTFLFAGFSTIGLF 539
 QY 492 FIYFVPEPKQSLAEIDQ 510

Db 540 FIWLLVPETKGLQFEVEK 558

RESULT 7
 G86812
 D-xylose proton-symporter [imported] - Lactococcus lactis subsp. lactis (strain IL14.
 C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 31-Mar-2001
 C:Accession: G86812
 R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malmgren, K.; Weissbach, J.; E.
 Genome Res. in press, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium.
 A:Reference number: A86625
 A:Accession: G86812
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-433 <STO>
 A:Cross-references: GB:AE005176; NID:g12724500; PID:AAK05601.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: xylT
 C:Superfamily: glucose transport protein

Query Match 18.98; Score 521.5; DB 2; Length 433;
 Best Local Similarity 28.98; Pred. No. 1.9e-27;
 Matches 146; Conservative 79; Mismatches 158; Indels 123; Gaps 11;
 QY 17 LGGLTFGYELAVISGALLPLQLDFGLSCLEQ-----FLVGSLLGLALLSLV 64
 Db 15 LGGLTFGYDTGVISGALL-----FIEKESWQVSSWAMMEGWTAAVLMGAVIGAV 65
 QY 65 GGLFDICYRKQAILGSLNVLGSLTLGLAGSLAWLVLGKAVVGFPAISLSSMACCIYVS 124
 Db 66 IGPMSDRFCRKRLLLLSAVIFFVGLGSLGSLNSAEILLISRVILGKAVGASALVPTLYLS 125
 QY 125 ELVGPQRGVLSLYEAGITVGLISYALNYALAGTPWGRHMFQWATAPVQLSLSLIF 184
 Db 126 ELSPAKIRGVSTMTFQMLMTGILLAYISNYALKVSGNWHMMLGLATVPVALLFIGGLF 185
 QY 185 LPAGTDEATHKDLIPLOGEAPKLG-----PGRPRYSFLDL-FRARDNMRGR----- 231
 Db 186 LPESPRFLVRHDN---EAGAREILGNDDPNSTEAEISDQLMAKEKOGGLQELFQ 241
 QY 232 -----TTVGLGLVLPQGLVFOQLGQPNVLCYASTIFSSVGFHGGSAVLASVGLGAVKVAATLT 286
 Db 242 MSRPVLIAMGIAIFQVVMGNTVLYFAPSFVAVGF-GASAAALLAHIGIGIFNVIVTYI 300
 QY 287 AMGLVDRAGRALLAGCALMALSVSGILVSFAVPMDSGSCSLAVPNATGOTGLPGDSG 346
 Db 301 AMRYMDKYNRRWMLNFGWGM-----GISLVMVSGM-----ILAENA----- 338
 QY 347 LQDSSSLPIPTNEDQREPILSTAKTKPHRSGDPSAPPLALSSALPGPPLPARGHA 406
 Db 339 -----HIGFG----- 343
 QY 407 LLRWALLCLMVFSVAFSGFGPVTWLVSEIYVPEIRGRAFAFNSFWAANLFIISLF 466
 Db 344 --KYLAVIALTYITAFFSATWGPVMMYMGESFPKIRGLGNSFGAANVNAANVVSUTL 401
 QY 467 LDLTIGTIGLSWTFLLYGLTAVLGLGF 492
 Db 402 LPLLSFEFGTKGIFLIYAACCFLSIWF 427

RESULT 8
 T38125
 Myo-inositol transporter 2 - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T38125
 R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, May 1997

A:Accession: Z21772

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-557 <BAD>

A:Cross-references: EMBL:Z95334; PIDN:CAB08597.1; GSPDB:GN00066; SPDB:SPAC20G8.03

A:Experimental source: strain 972h; cosmid c20G8

C:Genetics:

A:Gene: SPDB:SPAC20G8.03

A:Map position: 1

C:Superfamily: maltose transport protein MAL61

Query Match 18.8%; Score 520.5; DB 2; Length 557;
Best Local Similarity 27.1%; Pred. No. 2.8e-27;
Matches 144; Conservative 90; Mismatches 189; Indels 109; Gaps 10;

QY 12 ASVSLGLTFCYELAVISGALLPLQLDPG--LSCLEQEFVLSLLGALLASLVGGFLI 69
DB 85 SAVAGISGLLFGYGVISGALVGLSGDLGHLVSSQKRELITSATSAALISATTSGWLA 144
QY 70 DCYGRKQATLGSNLVLLAGSLTGLAGSLAWLVGRVAVGFALISLSSMACCIYVSELVGP 129
DB 145 DWGKRLLLCADAIFVIGSVIIMASRNVAHVGRFVIGYIGLTSLVIPWYITELAPA 204
QY 130 RORGVLVSYEAGITVIGLLSYALNYALAGTGWHRHMGWATAPAVLQSLSLFLPAGT 189
DB 205 RLURGLVTVYVFTGGQLIAYSLNAAFEHVHGWRIMEFGIAGAPALGQLISLFWTPESP 264
QY 190 DETATH---KDLPILOGGEAPKLGPCRPRY-----SFLDLF 222
DB 265 RYLLRHNVKVKYKILSRTHPEAKPAETIAIKVLSIQGVKYVDFPEGNKFOHFFHSLKVL 324
QY 223 RARDNMGRRTTVGLGLVFOQLTGOPNVLCYASTIPSSVFGHGGSSAVLASVAGVAKVA 282
DB 325 TVPSNRRS-LFVIGCFLOWFOQSGTNAIOYFSAIFQSVGF---KNSISVIVGATNFV 380
QY 283 ATFLAMGLVDRAGRALLAGCALMALSVSGIGLVSAFVPMDSGSPCLAVPNATGOTGLP 342
DB 381 FTIVAFMFIDIRGRIRLLCTSAVM---IAGLALCAIAYHF-----LP 420
QY 343 GDSGLLQDSSLPIPTNEDOREPILSTAKTKPHRSGDPSAPPLALSSALPGPPLPA 402
DB 421 AD-----TQNTNSG----- 430
QY 403 RGHALLRWTTALLCLMVFSAPSGFGPVTWLVLSLEYVEIRGRAPAFCSNFNMAANLFI 462
DB 431 -----WQYVVLASIIIFLASYASGIGNIPW-QQAEILFPMVEVRALGAGFTAINWVGNLII 484
QY 463 SLUSFLDLICTIGLSWTFLLYGLTAVLGLGFIYLFVETKQSLAEIDQFOFK 514
DB 485 SASFLTMESITPTGTFTALFAGCFVGLVTSYFTYPELAGMSIENHKULEK 536

RESULT 9

A:Accession: A85433

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-493 <STO>

A:Cross-references: GB:NC_001269; NID:g7270615; PIDN:CAB80333.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4g36670

A:Map position: 4

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring

Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488

A:Accession: A85433

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-493 <STO>

A:Cross-references: GB:NC_001269; NID:g7270615; PIDN:CAB80333.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4g36670

A:Map position: 4

C:Superfamily: glucose transport protein

Query Match 18.4%; Score 508; DB 2; Length 493;

Best Local Similarity 25.5%; Pred. No. 1.7e-26;

Matches 140; Conservative 83; Mismatches 194; Indels 132; Gaps 9;

QY 11 CASVSLGLTFCYELAVISGALLPLQLDPGLSCLSOEFLVGLSLLGALLASLVGGFLID 70
DB 20 CAIVASIVSIIFGYDTGVMSGANVFIEEDLKTNDVQTEVLTLNLNLCALVGSLLAGTSD 79
QY 71 CYGRKQATLGSNLVLLAGSLTGLAGSLAWLVGRVAVGFALISLSSMACCIYVSELVGP 130
DB 80 IIGRRITIVLASILFPMGLSILMGNGPNYPVLLSGRTAGLVGFALMVAVPYSAELATAS 139
QY 131 ORGVLSVLEAGITVIGLLSYALNYALAGTTPW--GHRHMGWATAPAVLQSLSLFLPAG 188
DB 140 HRGLASLPHLCISIGLLGYIVNPFSLPMHIGWRLMLGIAAIVPSLVLAFLGILKMPES 199
QY 189 -----TDATATHKDLPILOGGEAPKLGPCRPRYSFLDLPRARD-----NMRGT 232
DB 200 PRLIMQGRLEKEKEILELVSN-----SPEAELEFQDIKAAAGIDPKCYDDVVKMEGKK 254
QY 233 TVG-----LGLVLFQOLTGOPNVLCYASTIPSSVFGHGGSSAV 270
DB 255 THGEGVWKEILRPDPVRRVLLTALCIHFQHASGIEAVLLYGPRIFFKAGITTKDKLF 314
QY 271 LASVGLGAVKVAATLTMGLVDRAGRALL--AGCALMALSVSGIGLVSAFVPMDSGSPS 328
DB 315 LVTVIGVIMKTFITATLLLDKVRKRLTLTSVGNVIALTMLGFLT----- 363
QY 329 CLAVPNATGOTGLPGDSSLQDSSLPIPTNEDOREPILSTAKTKPHRSGDPSAPPR 388
DB 364 --MAQNAGGK----- 371
QY 389 LALSSALPGPPLPARGHALLRWTTALLCLMV---FVSAPSGFGPVTWLVLSLEYVPEIRG 445
DB 372 -----LAWALVLSIYAAYSFAFFSIGLIGPITWVYSEVFPKLRA 412
QY 446 RAPAFCSNFNMAANLFIISLFLDLICTIGLSWTFLLYGLTAVLGLGFIYLFVETKQSL 505
DB 413 QGASLGAVAVNRVNVATVNSHSLTSLSAITTGGGAFFMAGVAAVAMNFFPFLPETKQSL 472
QY 506 AEIDQFOFK 514
DB 473 EEIEALFOR 481

RESULT 10

F65079

galactose-proton symport (galactose transporter) - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 24-Sep-1999

C:Accession: F65079

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.: Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: F65079

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-464 <BLAT>

A:Cross-references: GB:AE0000377; GB:U00096; NID:g2367178; PIDN:AACT5980.1; PID:g17893

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: galP

C:Superfamily: glucose transport protein

Query Match 18.0%; Score 499; DB 2; Length 464;

Best Local Similarity 27.2%; Pred. No. 6.2e-26;

Matches 141; Conservative 95; Mismatches 183; Indels 100; Gaps 10;

QY 14 VSLGGLTFYGLAVISGALLPLQDLFGSLCEQELVGLSLLGALLASLVGGFLIDCYG 73
Db 21 LAALAGLLGLDVGIAAGLPFIADQITSHTOEWVSSMFWGAAGVAGVSGWLSFKLG 80
QY 74 RKQAILNSLVLLAGSUTLGLAGSLAVLVGRVAVFAISLSSMACCIYVSELVGRQRG 133
Db 81 RKSLMIGAILFVAGSLFSAAPNVEVLILSRVLGLAVGVASYTAPLYLSEIAPEKRG 140
QY 134 VLVSLEYAGITVGLLSYALNYALAGTPWGRHMFHGNATAPAVLQSLSLFLPLAGTDETA 193
Db 141 SMISYQIMITIGILGAYLSDTAFSYT-GAWRMWMLVIIIPAILLLIGVFFLPDPSRWFA 199
QY 194 THKDLIPLO-----GGEAPK-LGPRP-----RYSELDLPRARDNRGRITVGLGLV 239
Db 200 AKREVDVAERVLLRLRDTSAEAKRELDEIRSLQVKGSGWALFENSFRRAVFLGLVQ 259
QY 240 LFQOLTQOPNVLVASTFSSVGFHGGSSAVLASVGLGAVKVAATLTAMGLVDRAGRAL 299
Db 260 VMOQFTGMNVMYAPKIFELAGVTNTTEQMWGTIVVGLTNVLATFTAIGLVDWRGKPT 319
QY 300 LLACALMALSVSGIGLVSAFVPMDSGLAVPNATGOTGLPGDGLDSSLPPIPRPT 359
Db 320 LTGLFLVNA---AGMGVIG-----
QY 360 NEDOREPILSTAKTKPHPRSGDPSAPPRIALSSALPGPPLPARGHALLRWTALLCLMVF 419
Db 336 -----TMHGIHSPSA-----QYFAIALLMFP 358
QY 420 VSASFSGFPGVTWLVLSIYVPEIRGAFAP-CN-SFNWAANLFISSLFDLIGTGLSW 477
Db 359 IVGFAMSAAGPLIWLVCSEIQP--LKGRDFGTCSTATNLIANNIVGATFTMLTILGNAN 416
QY 478 TFLYGLTAVLGIGFYVFPETKQSLAEIDQOFKRR 516
Db 417 TFWYAAALNVFLILLTLWLVPETKHYSLHIERNLKGR 455

RESULT 11
A:Accession: A48442
A:Title: membrane transport protein (clone DL.SH) - Leishmania donovani
C:Species: Leishmania donovani
C:Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999
C:Accession: A48442
R:Langford, C.K.; Ewbank, S.A.; Hanson, S.S.; Ullman, B.; Landfear, S.M.
Mol. Biochem. Parasitol. 55, 51-64, 1992
A:Title: Molecular characterization of two genes encoding members of the glucose transp
A:Reference number: A48442; MUID:93063058
A:Accession: A48442
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-547 <LAN>
A:Cross-references: GB:M85072; NID:g159301; PIDN:AAA29230.1; PID:g159302
A:Note: sequence extracted from NCBI backbone (NCBIN:118159, NCBIP:118162)
C:Superfamily: glucose transport protein
C:Keywords: transmembrane protein

Query Match .18.0%; Score 497.5; DB 2; Length 547;
Best Local Similarity 26.4%; Pred. No. 9.4e-26;
Matches 146; Conservative 83; Mismatches 190; Indels 135; Gaps 15;

QY 10 LCASVSLGLLTFGYELAVISGALLPLQDLFGSL--CLEQELVGLSLLGALLASLVGGF 67
Db 7 LCAN---LGGFLFGYDTGVINIAALFQMDHFGFSEHSQYALIVAIAIAGAFVAFISGF 63
QY 68 LIDYGRKQOAILGSLNVLVLAGSLTLGLAGSLAWLVGRVAVGFATLSLSSMACCIYVSELV 127
Db 64 ISAAFGRRPCIAVADALFVIGSVLMGAAPNVEVLVSRVIGLVGLIGSATTPVYLAET 123
QY 128 GPRQGVLSLYEAGITVGLLS---YALNYALAGTPWGRHMFHGNATAPAVLQSLSLIF 184
Db 124 SPKRGATIVLNNLPLTGTGQFVAAGFTAIMVFTSKNIGRWRAIGIGALPAVYQAFCLLF 183

QY 185 -----LPAGTDETTAT-----HKDLIPLO--GGEAPKLGPRPRYSELDLFRARDNM 228
Db 184 FLPESPRWLLSKGHADRAKAVADRFEVDLCFEQFGDELPSV-----RIDYRPLM-ARD-M 236
QY 229 RGRITVGLGLVLFQOLTQOPNVLVASTFSSVGFHGGSSAVLASVGLGAVKVAATLTAM 288
Db 237 RFRVVLSSGLOITQOFGSINTIMYSSVILYDAGFRAIMPVVLSIPLAFMNALFTVAI 296
QY 289 GLVDRAGRARRALLA---GCALMALSVSGIGL-----VSFAVPMDSGLAVPNATGOTG 340
Db 297 FTVDREGRERRMLLSVFGCLVLLVIAIIGFFIGTRISYSV----- 337
QY 341 LPDGSGLLDSSLPPIPRPTNEDOREPILSTAKTKPHPRSGDPSAPPRIALSSALPGPPL 400
Db 338 ---GGGLF----- 342
QY 401 PARGHALLRWTALLCLMVFSAFSGFPGVTWLVLSIYVPEIRGAFAPCNSEFNWAANL 460
Db 343 -----LALAVFLALYAPGICIPWIMGIEIFPTHLRTSAAVSATMANWGANV 390
QY 461 FISLFDLIGTIGLSTWTLFLLYGLTAVLGIGFYVFPETKQSLAEIDQOFKRR----- 515
Db 391 LVSOVFPILMGAIGVGCTFTIISLMALGICFYVFFFAVETKGLTLEIDNMFRRAGLPP 450
QY 516 RFTLSFGHRQNSTG 529
Db 451 RF-----HEEGSG 459

RESULT 12
A:Accession: A84537

Probable sugar transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: A84537
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487

A:Accession: A84537

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-511 <STO>

A:Cross-references: GB:AE002093; NID:g4678209; PIDN:AAD26955.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g16130

A:Map position: 2

C:Superfamily: glucose transport protein

Query Match .18.0%; Score 497; DB 2; Length 511;
Best Local Similarity 26.5%; Pred. No. 9.4e-26;
Matches 149; Conservative 80; Mismatches 197; Indels 136; Gaps 11;

QY 11 CASVSLGLLTFGYELAVISGALLPLQDLFGSLCEQELVGLSLLGALLASLVGGFLID 70
Db 29 CAILASMTSIIIGYDVGWVGAAIFIKDKLUSDVOLEILMGLINLIYSIGSAAAGRTSD 88
QY 71 CYGRKQOAILGSLNVLVLAGSLTLGLAGSLAWLVGRVAVGFATLSLSSMACCIYVSELVGR 130
Db 89 WIGRRYTVLAGTFFFCGALLGFATNYPFINVGRFVAGVGYAMTAPVYTTVEAPAS 148
QY 131 QRGVLVLSYEAGITVGLLSYALNYALAGTP--WGRHMFHGNATAPAVLQSLSLFLPAG 188
Db 149 SRGFLSSPPEIFINIGILLGYVSNYFFAKLPEHIGWRMLGIGAVSPVFLAIGVLAMPES 208
QY 189 -----TDETATHKD-----LIPLOGEAPK 208
Db 209 PRWLWQGRGLGDAFKVLDTKTSNTKEAISRLNDIKRANGIPDDMTDDVIVVP-----NK 262

Db 309 SGISNNHTTIWISVLLSLCNPIGFVPMSLIEKVGRRRIIFLSCGLVLSLVFVIGVAFLL 368

Qy 321 VPMDSGPS-----CLAVPN-----ATGQTGLPGDGLQDSSLPPI 356

Db 369 VNHDSAAATLPANQVSNFNSSYPDAKCMAYSNCDYCVTTDACGFCCHDANTKQGYCLP-- 426

Qy 357 PRYNEDOREPILSTAKTKPKHPSGDSAPPRIALSSALPGPPLPARGHALLRWTAALLCL 416

Db 427 --AGFDNPEVYSSTGCTNSNGSTANNFKWEKYCYDPTYLLPIIACG----- 472

Qy 417 MVVSAFSFGPGPVTVLVLSEIYVPEIRGRAFCPNSFNNAANLFISSLFDLDTGTTGLS 476

Db 473 -VYLTTSSGFTSLPWNLSSEFYPMWARSTCVALTSTSNWVFNIIALTLSLTQVIGKY 531

Qy 477 WTFLLYGLTAVLGLGFIYLFVETPKGOSLAEIDQOF-----OKRRPTLS-FGHRQ 525

Db 532 GAFWLVAGLTVIAPIFILFLVETPKGYSIEEVENLPMKKORREASRRRETIVTEVSRM 591

Qy 526 NST 528

Db 592 NST 594

RESULT 14

F85951

galactose-proton symport of transport system [imported] - Escherichia coli (strain O1

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001

C:Accession: F85951

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grotbeck, E.O.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potemousis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: F85951

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-464 <STO>

A:Cross-references: GB:A0051174; MID:g12517486; PIDN:AAG58074.1; GSPDB:GN00145; UWGP:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: galP

C:Superfamily: glucose transport protein

Query Match 17.9% Score 496; DB 2; Length 464;

Best Local Similarity 27.2%; Pred. No. 9.9e-26;

Matches 141; Conservative 95; Mismatches 183; Indels 100; Gaps 10;

Qy 14 VSLGLGTFGYELAVISGALLPLQDLFGSLCEQEFVLGSLILGALLASLVGGFLIDCYG 73

Db 21 LAALAGLLFGLDIGVAGALPFIADFEQITTSHTQEWVSSMMFGAAGVAGSGWLSFKLG 80

Qy 74 RKQAILGSLNVLGASITLGLAGSLAWLVGRAVVGFAISLSSMACCIYVSELVGPQRG 133

Db 81 RKSLMIGAILFVAGLSFSAAPNVEVILSRVLLGLAVGVASVYAPLYLSEIAPEKIRG 140

Qy 134 VLVSLEYAGITVGTLLSYALNYALAGTPWGRHMFGNATAPAVLOSLLPLPAGTDETA 193

Db 141 SMISWYQMITIGLIGAYLSDTAFSYT-GAWRMWLVIIIPAILLLIGVFLPDPSPRMFA 199

Qy 194 THKDLIPLQ-----GEAPK-LGPGRP-----RYSFDLDFRARNMRGRTTVGLGLV 239

Db 200 AKRFVDAERVLLRLRDTSAEKRELDEIRSLQVKSGHWALFKENSNFRVFLGVLLQ 259

Qy 240 LFOQLTGQPNVLCYASTIFSSVGHGSSAVLASVGLGAVKVAATLTAMGLVDRAGRRL 299

Db 260 VMOQFTGMNVMYAPKIFELAGYNTTQEWGQTVIVGLTNVLTATFIAIGLVDRWGRKPT 319

Qy 300 LLACCALMALSVSGIGLVSAFVPMVDPSPCLAVPNATGQTGLPGDGLQDSSLPPIPT 359

Db 320 LTLGLFLVNA---AGMGVLG----- 395

QY 360 NEDQREPISTAKTKPHRPSGDPSPAPRLALSSALPGPPLPARGHALLRWTAALLCLMV 419
Db 336 -----TWMHGHSPPA-----QYFAIAMLLMF 358
QY 420 VSAFSGFGPVTWLVSEIYVPEIRGRAPAF-CN-SFWAANLFISSLFSLDGLIGTIGLSW 477
Db 359 IYGFANSAGPLIWLVCSEIQP--LGRDRFGITCSTATNANNIVGATFLTMLNLGNAN 416
QY 478 TFLLYGLTAVLGLGYLYFVPEVKGSGLAEIDQOQOKR 516
Db 417 TFWVYAGLVNLTLLVLPVPEVKHSHLEHIERNLKGR 455
RESULT 15
B26430
L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli
C:Species: Escherichia coli
C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 24-Sep-1999
C:Accession: B26430; A28075; I40996; B65067
R:Maizden, M.C.J.; Davis, E.O.; Baldwin, S.A.; Moore, D.C.M.; Henderson, P.J.F.
Nature 325, 641-643, 1987
A:Title: Mammalian and bacterial sugar transport proteins are homologous.
A:Reference number: A93389; MUID:87115869
A:Accession: B26430
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-472 <MAI>
R:Maizden, M.C.J.; Jones-Mortimer, M.C.; Henderson, P.J.F.
J. Biol. Chem. 263, 8003-8010, 1988
A:Title: The cloning, DNA sequence, and overexpression of the gene araE coding for arabinose isomerase.
A:Reference number: A28075; MUID:88228015
A:Accession: A28075
A:Molecule type: DNA
A:Residues: 1-472 <MA2>
A:Cross-references: GB:J03732; NID:gl45320; PIDN:AAA23469.1; PID:gl45321
R:Stoner, C.; Schleif, R.
J. Mol. Biol. 171, 369-381, 1993
A:Title: The araE low affinity L-arabinose transport promoter. Cloning, sequence, transcription, and overexpression.
A:Reference number: I40996; MUID:84114868
A:Accession: I40996
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-25, 'Y', 349, 'R' <RES>
A:Cross-references: EMBL:X00272; NID:g40940; PIDN:CAA25075.1; PID:g40941
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, D.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: B65067
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-472 <BLAT>
A:Cross-references: GB:AE000368; GB:U00096; NID:g2367165; PIDN:AAC75880.1; PID:g1789207
C:Genetics:
A:Gene: araE
A:Map position: 61 min
C:Superfamily: glucose transport protein
C:Keywords: arabinose transport; intramolecular oxidoreductase; isomerase; membrane protein

Query Match 17.9%; Score 495; DB 2; Length 472;
Best Local Similarity 25.2%; Pred. No. 1.2e-25;
Matches 137; Conservative 93; Mismatches 182; Indels 132; Gaps 8;

QY 17 LGGLTFGYELAVISGALLPLQIDFGLSCLEQFELVGLSLLGALLASLVGGFLIDCYGRKO 76
Db 31 VAGLEFLGIDIGVAGALPFITDHFVLTSLRQEWVSSMMLGAAIGALFNGLSFLGRKY 90
QY 77 AILGSNLVLLAGSLTGLIAGSLAVLGRVAVGFATLSMMACCIIYVSELVGPQRQGLV 136
Db 91 SLMAGAILFVLGSGAFATSVEMLIAARVVLGAVGIASTAPLYLSEMASENVRCKMI 150

QY 137 SLYEAGITVGILLSYALNYALAGTPGWRHMGWATAPAVLQSLSLFLP----- 186
Db 151 SMYQLMVTGLGIVLAFSLDTAFSYS-GNWRMLGVLAIPAVLLIILVVFSPRMLAEKG 209
QY 187 -----AGTDETA-----THKDLIPLQGGAPKLGPRPRYSFUDLFRADNM 228
Db 210 RHIEAEVLRMLRDTSEKAREELNEISLKKOGWA-----LFIKINRV 255
QY 229 RGRITVGLGLVLFQOGLGOPNVLCYASTIFSSVGHGSSAVLASVGLGAVKVAATITAM 288
Db 256 BRAVFLGMLLQAMQOFTGMNIIYAPRIKMAAGFTTTEQQMIATLVVGLTFMFATEIAV 315
QY 289 GLVDRAGRRALLLAGCALMALS--VSGIGLVSEAVPMDSPSCLAVPNATGOTGLPCDSG 346
Db 316 FTVDKAGRKPAKLTGFSVMALGTLVLGYCLMQF-----DNGTASSG----- 356
QY 347 LLQDSSLPPIPRTNEDQREPISTAKTKPHRSGDPSAPRLALSSALPGPPLPARGHA 406
Db 357 ----- 356
QY 407 LLRWTALLCLMVFSAFSGFGPVTWLVSEIYVPEIRGRAFAFCNSFNAAANLFISSLF 466
Db 357 -LSWLSVGMVTMMCIAGVMSAAPVWILCSEIQPLKCRDFGTCSTTTNWSNIIIGATF 415
QY 467 LDIGRIGLSWTFLYGLTAVLGLGYLYFVPEVKGSGLAEIDQOQOKRRFTLSFGHRQN 526
Db 416 LTLLDSIGAAGTFWLYTALNIAFVGITFWLIPETKNVTLIEHIERK-----LMAGEKLR 468
QY 527 STGI 530
Db 469 NIGV 472

Search completed: March 15, 2002, 07:07:47
Job time: 1463 sec